

複数の観点で分類した自然言語処理用シソーラス

国分 芳宏[†]・岡野 弘行[†]

従来の情報検索に特化されたシソーラスではなく、構文解析や用語標準化などの自然言語処理を目的とする 420,000 語規模のシソーラスを開発した。各用語の持つ関係語の数が膨大なため、観点（ファセット）を導入して分類し、探しやすいシソーラスである。さらに、差別語、表記の揺れなども区別できる。シソーラスを作成する際の留意点・課題もまとめた。パッケージソフトのカスタマイズ機能およびインターネットや他の辞書との連動機能、用語の標準化などについても紹介した。

キーワード：シソーラス、観点、自然言語処理、用語標準化

A Thesaurus which Classifies Terms by Facets for Natural Language Processing of Japanese

YOSHIHIRO KOKUBU[†] and HIROYUKI OKANO[†]

Instead of a thesaurus specialized in conventional information retrieval, we developed a thesaurus of 420,000 terms for the purpose of natural language processing such as parsing or the term standardization. Because each entry term has a large number of terms with various semantic relations, we introduce a facet and classify them for finding relative terms easily. Furthermore, we distinguish discriminatory terms, and fluctuating Japanese spellings. We described points to keep in mind and future tasks in making a thesaurus. Our package has the connecting function with the Internet and the other dictionaries.

Key Words: *Thesaurus, Facet, Natural Language Processing of Japanese, Term standardization*

1 はじめに

筆者らは、1990 年、自然言語処理のための解析辞書の日本語表記の揺れを管理することから始め、1995 年に同義語辞書の初版を発行した。その後、用語の意味関係を含むシソーラスのパッケージを発売し、現在 6 版を重ねている。1 年間に 20,000 語程度を追加していて、420,000 語に達している。

これまでのシソーラスは、主として、情報検索のキーワードを選択するための支援ツールとして開発されてきた。登録されている用語は該当する分野の専門用語が主体で、さらに品詞は名詞だけであった。そのため、情報検索を越えて、文書整理や統計処理などのために必要な構

[†] 株式会社言語工学研究所, Institute of Language Engineering

文解析や用語の標準化など、自然言語処理に利用することは難しかった。

筆者らのシソーラスは、自然言語処理を目的とした一般語を主とするシソーラスである。いわゆる名詞だけでなく、動詞、形容詞、形容動詞、副詞、代名詞、擬態語さらに慣用句までを登録している。これまでのシソーラスでは、作成者の考え方で分類してあった。使用者は、その分類基準に従ってたどって探さなければならなかった。また紙面の物理的な制約もあって意味空間を1次元的に整理してあった。本来意味分類は多次元空間のはずで、筆者らのシソーラスでは、複数の観点で多次元的に分類してある。

また、メール整理に代表されるような文書整理のために、時事的な用語や省略語も積極的に登録している。送り仮名や訳語などの差異による異表記語も網羅的に収集した。自然言語処理で使うことを目的としているため、テレビなどから収集した新語や、構文解析で発見した新語を登録している。

用語間の意味関係として、広義—狭義（上位—下位）関係、関連関係および同義—反義関係を持っている。

流動的に変化する用語の意味および用語間関係への対応とコスト・パフォーマンスの観点から、トップダウン方式ではなく、ボトムアップ方式で開発した。一般語を主体としているが、他の専門シソーラスと併合もできる。

以下、（第2章）用語の収集とシソーラスの構造、（第3章）用語同士の意味関係、（第4章）パッケージソフトの機能について順次述べる。

2 用語の収集とシソーラスの構造

用語の収集と分類の仕方について述べる。

2.1 用語の収集

市販の辞書は語義が分からない用語を調べるためのものである。筆者らのシソーラスは自然言語処理で使うのが目的なので、市販の辞書に記載されている用語よりも、頻繁に使われる用語を中心に登録している。正しい表記だけでなくよく使われるのであれば「キューピット」(Cupid)のように誤った表記も登録している。

筆者らの構文解析（別途 SDK として販売）で新しい記事コーパスを解析した結果、解析できなかった新しい用語は逐次、解析辞書に登録しているが、同時にシソーラスにも登録している。シソーラス更新中に、追加登録した用語も形態素解析の辞書に登録している。見つけた用語がよく使われている用語かどうかはネットで調べている。

新語を探しだす作業よりもシソーラス上の用語と関連付ける作業の方が工数がかかった。

品詞分類も名詞の意味を除いて解析辞書と同じにしてある（文末付録参照）。構文解析では名

詞を意味で分類しているが、シソーラスでは意味分類が構文解析より詳細なこと、名詞に複数の意味を持たせられないことなどの理由で意味では分類していない。

自然言語処理用のシソーラスではほとんど全ての自立語が収集の対象になる。構文解析では、解析時に結合して処理するので、複合語の要素だけを網羅すれば十分である。一方シソーラスでは組み合わせられた複合語もすべて網羅する必要があるため、語数が多くなる。

自然言語処理では固有名詞も重要な位置をしめるが、日々生成されていて、かつ変化するため辞書として登録するには手が付けられないので、地名など変化の少ないもの以外は登録できていない。

2.2 シソーラス構造における分類

複数の観点での分類

意味空間は1次元ではなく多次元である。どの属性に注目して（観点で）分類するかによって、いろいろな分類の仕方が考えられる。身近な例で「料理」について考えてみる。古今東西の料理の種類は相当な数になり、分類の仕方も人によって異なる。ここで調理法、材料、地域の3つの観点で分類するとつぎのようになる。

調理法の観点で分類すると 生もの、煮物、焼き物

材料の観点で分類すると 魚料理、肉料理、野菜料理

地域の観点で分類すると 和食、中華、洋食

例えば「刺し身」は、料理を3つの観点によって分類した結果、連想された用語「魚料理」「生もの」「和食」の狭義語である。逆に「刺し身」の広義語が「生もの」「魚料理」「和食」の3つあることになる。その結果、網構造になる。これを図にすると、図1のようになる。

この他に「料理」のための観点としては「対象」（病人食、独身料理）「スタイル」（会席料理、飲茶）などが考えられる。いろいろな考え方で探す利用者がいるので、なるべく多くの観点で分類しておく必要がある。

人間の感覚に沿った分類

色を分類するときにも参考文献にあげたシソーラスでは「赤系統」「青系統」「黄色系統」などと色相や明度などに従って分類してある。データベースの検索の支援をするためには、人間との関係を重視して「はでな色」「暖かい色」といった人間の感覚に沿った観点での分類も併設した方が実用的である。筆者らのシソーラスもなるべく多くの観点で分類している。

検索された用語を見やすくする目的で、グループに入れる用語を少なくする方針を取ったため階層が深くなってしまった。電子化されたシソーラスでは、クリックするだけで、簡単に上下の階層に移行できるので階層を深くしても問題は少ないのであるが、グループにつける名前が恣意的になりがちなのが課題である。

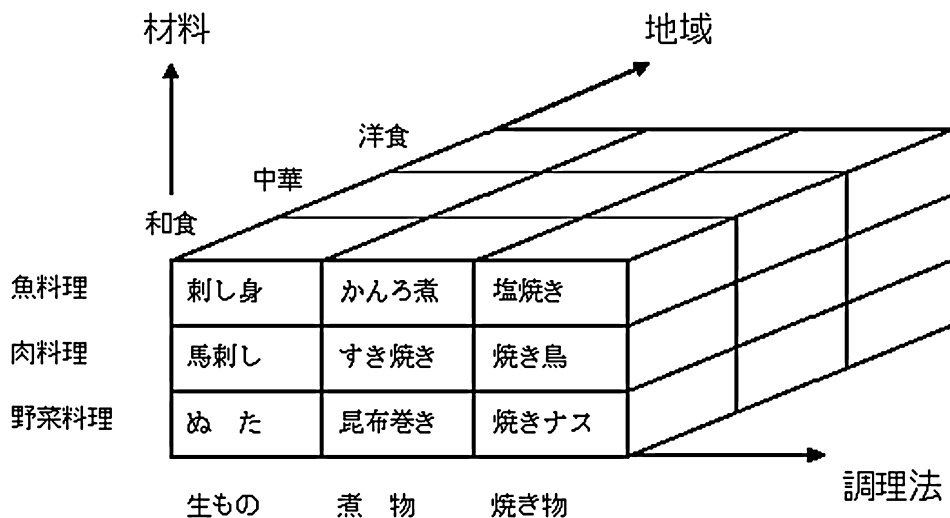


図 1 「料理」を「調理法」「材料」「地域」の 3 つの観点で分類した例

分類作業における揺れの吸収

用語同士の意味的な関係は、自明な場合だけではない。どこまでを同義語として認めるかは、シソーラスの作業員同士でも食い違うことがある。現在 3 名で相談しながら、最終的には多数決で決めている。

例えば「明日」と「翌日」を考えて見ると、意味的にほとんど重なっていて同義語と思われるが、厳密に言えば違いがある。同義語にするか関連語にするかが一義的に決定できない。

	「明日」	「翌日」
現在	○	○
過去	×	○

このように微妙に意味の異なる場合にユーザーの意見を聞いて、同義語として扱った場合が多い。

2.3 シソーラス構造における記述

補助的な記述

各の用語の持つ関係語の数が多いため、用語を 3 つの目的で、(|) で区切って補助的な記述をつけている。

A. 分類の観点を表示する。

狭義語の例

料理 材料	肉料理, 魚料理, 野菜料理
料理 地域	和食, 洋食, 中華料理

B. 同じ分類に属する用語が膨大な数になるため細分したいときに, 細分した分野に対応する適当な用語がなく, 恣意的な用語になるのを防ぐ.

狭義語の例

肉料理 | 煮物 シチュー
肉料理 | 薫製 ビーフジャーキー

C. 多義語を区別する.

狭義語の例

月 | 天体 満月, 寒月, 三日月
月 | 時間 正月, うるう月

現在関係語が多い用語を中心に, 全体の 4.0 パーセントの用語に補助的な記述が付けてある.

語義の違いの記述法

冊子体のシソーラスは木構造で, その構造をたどりながら探していかなければならなかった. 筆者らの電子化されたシソーラスはキーボードから直接構造上のどこでも指定できるので, もはや木構造である必要はない. 網構造で, 複数の広義語を持つことになる. しかしその結果同じ文字列で複数の意味を持つ多義語が区別できないという問題がでてくる. 例えば木構造で検索したときには, 「時間」からたどった「月」(month) と, 「天体」からたどった「月」(moon) の 2 つの異なった意味の用語は区別できるが, 直接「月」と指定する方法では区別ができなくなる.

補助的な記述をつけて「月」を「天体」の観点でとらえたときは「月 | 天体」で「時間」の観点でとらえたときは「月 | 時間」として区別した.

広義語	狭義語
月 天体	名月
月 時間	正月

用言

自然言語処理で使うには, 名詞だけでなく用言 (動詞, 形容詞) や副詞も登録しておく必要がある. 用言は語幹と活用形で登録してある. パッケージソフトでは終止形で表示する.

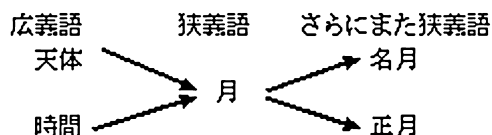


図 2 「月」を中心にした構造

活用形も構文解析に合わせてある.

例 動(く) 動詞 カ行 5 段活用

赤(い) 形容詞

慣用句

日本語では、慣用句が大きな意味的位置を占めている.

慣用句はまとめた形で 1 語にして登録してある.

例 「水をあける」 = 「引き離す」

「水をあける」は「引き離す」という意味で「水」の意味はまったくない. 「水をあけ(る)」は 1 つの動詞にして「引き離す」の同義語として登録してある.

慣用句は用法によって間に挟まれる助詞までが変わるものがある.

「山田は顔が広い」(叙述用法) 「顔が広(い)」を形容詞として登録してある.

「顔の広い山田は」(限定用法) 「顔の広い」を連体詞として登録してある.

誤りのある用語

実際にシソーラスを運用するためには、関係する用語として差別語を出力しないなどといった細かい配慮が必須である. 差別語は年々増える方向にある. 増える差別語を次々に登録していくためにもいつもシソーラスを更新していかなければならない.

エラーに 2 つのレベルがある.

誤り語および差別語.

誤り語 例 ご多聞にもれず (正: ご多分にもれず)

差別語

標準でない用語.

常用漢字以外を含んでいる用語

表記の揺れ 例 インタフェース (正: インターフェース)

旧地名 例 浦和市

旧機関名 例 文部省

商品名 例 宅急便

3 用語同士の意味関係

用語同士の意味関係として、表 1 のものを用意した. 広義語—狭義語の関係は自然言語処理で広義語に適用した規則は、狭義語にも適用できるようにするため狭義語になるのは同じ属性のものだけとした. 「自動車」—「タイヤ」のような「全体」—「部分」関係は「部分」という観点の関連語とした. 原則として自立語だけとしたが、一部に接尾辞も採択してある.

3.1 同義語

英語で 1 人称単数は「I」だけであるが, 日本語には「私」「僕」「我」「小生」「我が輩」「手前」「愚生」と数十あり, 話者と相手との関係で使い分けられている. 日本語にはなぜ同じ意味の用語, 同義語がこんなに多いのか考えてみる. (表 2 参照)

辞書の中では, 「大和言葉」「外来語」などの区別はせず, 同等に扱っている.

外来語

日本語のなかに奈良時代には中国, 朝鮮から, 最近は主に米国から輸入されて日本語の中に入ってきている用語がある. 多少のニュアンスの違いはあるが, すべて同義語といえる. このような組み合わせが日本語のなかにたくさんあり, これが同義語を増やしている大きな原因である. 大和言葉は親しみやすさを, 漢語は権威を, 片仮名語は近代的な感じをあたえる. また最近「計算機」が「コンピューター」に, 「写真機」が「カメラ」になるといったふうに, 漢語が片仮名語に置き換わる傾向がある.

通称

通称と正式名称が両方使われている.

「首相」 = 「内閣総理大臣」

表 1 用語同士の意味関係

同義語	例 「犬」から見た「ドッグ」 表記の揺れも含む.
反義語	例 「強い」から見た「弱い」
狭義語	原則として属性の同じもののみとした. 例 「犬」から見た「秋田犬」
広義語	例 「犬」から見た「哺乳類」
関連語	例 「犬」から見た「キツネ」 「犬」から見た「犬小屋」 品詞の異なる用語, 自動詞—他動詞の対応なども関連語とした.
共起語	係り受け関係を構成する用語の組. 係りと受けの間にはいる助詞および良しあしも管理している. 例 寿命 (が), 延びる 良い 望遠鏡 (で), 見る (パッケージソフトには含めていない.)

表 2 同義語の例

大和言葉	漢語 (複合語)	片仮名語	英 字
打ち合わせ	会議	ミーティング	
しお	食塩	ソルト	NaCl
	読み出し専用メモリー	ロム	ROM

年号

わが国だけの問題であるが、年号が2種類ある。さらに漢数字とアラビア数字が両方使われる。

「2008 年」＝「平成 20 年」＝「平成二十年」

立場による用語の違い

立場によって同じことを違った用語で表す場合がある。例えば「税金」という用語を政府は「公的資金」という言い方をするが、納税者は「血税」という言葉を使う。検索者は「税金」という用語で探すだろう。このような傾向は社会科学の用語に多い。

省略語

「特別急行」→「特急」のようなものをいうが、「マスコミ」は「マス・コミュニケーション」の省略形であったというように、現在は省略形の方が4拍の新しい用語として定着してしまっているものがたくさんある。省略の程度も地域によって異なる。関東よりも関西の方が積極的に省略するようである。

「弱冷房車」(JR 東日本)

「弱冷車」(JR 西日本)

頭字語（英語の用語の先頭の文字だけを集めた用語：アクロニム）もこの省略形に入れるべきだろう。

ROM Read Only Memory

表記の揺れ

同義語のうち発音も同じものを表記の揺れ（異表記語ともいう）と言う。日本語では標準とされている表記の他に複数の「表記の揺れ」が許されている用語がある。個人により、機関により、いろいろな表記が氾濫している。極端な場合には、同じ著者が書いた記事でも表記法が違ふことがある。複数の機関の記事を一度に検索しようとする場合には、考えられる揺れをすべてキーにして検索しなければならない。

漢字と仮名による表記の揺れ

犬, イヌ, いぬ

漢字表記の揺れ

沈殿, 沈澱 (「澱」の字が常用漢字でないので「殿」の字を代用した.)

超電導 (JIS)

超伝導 (学術用語)

外来語をカタカナ書きするときの揺れ

インターフェース (新聞 1996 年まではインタフェース)

インタフェース (JIS)

インターフェイス (学術用語)

インタフェイス

古い記事を扱うときは異体字も問題になる.

国語, 國語

送り仮名の違いによる表記の揺れ

行う, 行なう

打ち合わせ, 打ち合せ, 打合わせ, 打合せ, 打合

(内閣告示の「送り仮名の付け方」の中にも複数の表記が許容されている.)

推奨語

用語を標準化するために, 同義語のグループのなかから, 言語工学研究所が推奨する用語である. 標準の用語に置き換える機能はパッケージソフトには含まれていない. 別売の用語標準化ソフトとして提供している.

インターフェース (新聞)

インタフェース (JIS)

インターフェイス (学術用語)

インタフェイス

⇒ インターフェース (言語工学研究所推奨)

米, 米国, USA, U.S.A., アメリカ合衆国, 合衆国, アメリカ (新聞)

⇒ アメリカ (言語工学研究所推奨)

3.2 反義語

意味が対立する用語の関係である. 対立の仕方にいくつかある.

A. 片方を否定すると対立する相手になる用語の関係である. 「良いこと」を否定すると「悪いこと」になるような関係である.

例 善 \longleftrightarrow 悪

B. ある中間的な点を中心にして逆の方向になる用語の関係である.

例 上 \leftarrow 中 \rightarrow 下

C. 一つの行為を対立する立場で捕らえた用語の関係である.

例 売る \longleftrightarrow 買う

D. さらには「兄」に年齢で対立する用語として「弟」がある. また性別で対立する用語として「姉」がある. どちらも反義語になる.

例 兄 \leftarrow 年齢的対立 \rightarrow 弟

↑

性別的対立

↓

姉

3.3 広義語・狭義語

自然言語処理で広義語との関係が狭義語にも適用できるように広義語・狭義語の関係は、属性が同じものだけにした。「自動車」―「タイヤ」のような全体部分関係は関連語にした。

- 例 1 東京都 新宿区 (狭義語)
東京都 都庁 (関連語)

「東京都に住む」「新宿区に住む」は成り立つが、「都庁に住む」は成り立たない。

- 例 2 疾病 伝染病 (狭義語)
疾病 発病 (関連語)

3.4 関連語

ある程度の意味的な関連性を持つ用語の関係を言う。大きく分けると同じカテゴリーの用語と異なるカテゴリーの用語との関係がある。

A. 共通の広義語を持つ用語。

広義語 狭義語

食材 → 肉

↘

野菜

(「肉」と「野菜」とは関連語である。)

B. 異なるカテゴリーであるが、意味的な関係のある用語。

広義語 狭義語

食材 → 肉

料理 → 肉料理

(「肉」と「肉料理」とは関連語と、人の判断で設定する。)

3.5 多義語

英語は多義語が多いと言われているが、日本語、特に大和言葉も多義語が多い。関係語は別の言葉になる。

大和言葉での例

- うめる 穴をうめる。
お風呂をうめる。
借金をうめる。
時間をうめる。

外来語での例 英語の多義性の影響も受けている。

- ライト 光, 照明, 明るい, 軽い

右, 右翼手
権利
書く

多義語は, | で区切って補助的な記述を付けてそれぞれ別の語として扱っている.

3.6 共起語

係り受けを構成する組み合わせを集めた辞書である. 構文解析で係り先を決定したり, 「良しあし」を決定したりするときに用いる. 係り側の格助詞を含めて管理している.

構文解析の通常の単語には必要に応じて「良しあし」のフラグが振っており, リスク管理やリコメンデーションなどに使っている. しかし例えば下記の例では,

例 寿命 が 延びる (良い)

例 寿命 が 短い (悪い)

「寿命」, 「延びる」, 「短い」など用語はそれ自体では「良しあし」の情報は持っていないが, 係り受けになったときに「良しあし」の性質が出てくる.

共起語として, 7万組を登録してある. 係り, 受けのそれぞれの用語の同義語, 狭義語を実行時に拡張するので, 共起語辞書に登録されていない係り受けにも対応できる.

ビール が 冷えている.

麦酒 が 冷えている (同義語)

生ビール が 冷えている (狭義語)

組み合わせの意味的な関係として, 現在は「良い」「悪い」「それ以外」の3種類の情報しか持っていない. 将来オントロジーとして発展させていく予定である.

共起辞書の作成方法

コーパスを構文解析で係り受けファイルにして, その結果を整理して, 共起語辞書に登録する方式を取っている.

3.7 用語の意味関係は時と共に変化する

1991年当時「発泡酒」は「ビール」の意味を含む広義語であった (JIS X 0901-1991). しかし現在は「発泡酒」は「ビール」とは別のものの名前になったため, 「ビール」「発泡酒」は共に「醸造酒」の狭義語となった. シソーラスには, 原則として最新の意味だけを採択している.

広義語 狭義語
醸造酒 → ビール
 ↘
 発泡酒

4 パッケージソフトの機能

筆者らのシソーラスはパッケージで販売するほかにネットからも使えるようにしてある。

4.1 操作画面

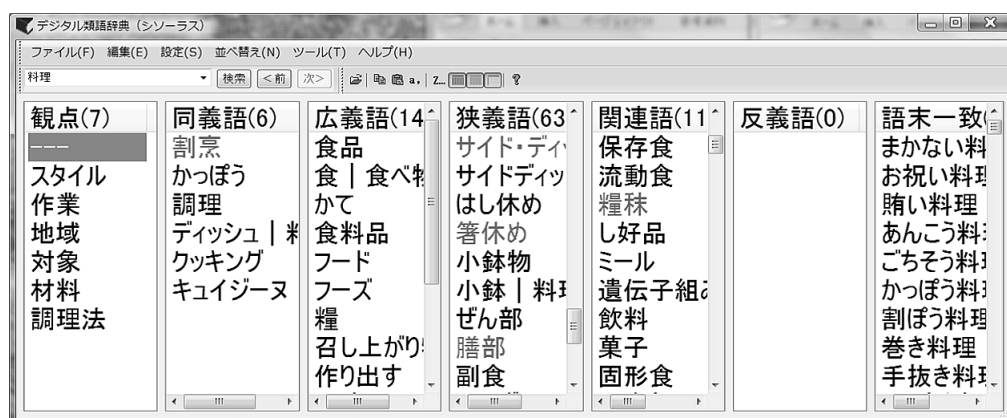


図 3 「料理」で検索したときの操作画面

観点の「——」は下に表示したどの観点にも属さない関係語を集めたものである。

「かつぼう」に対する「割烹」, 「はし休め」に対する「箸休め」など、標準でない標記も表示されている。実際の画面上では文字色で区別している。

(注)「エラーレベル・差別語の表示」(4.5)を参照のこと

以下に「料理」を「調理法」「材料」「地域」の観点で検索したときの画面を示す。

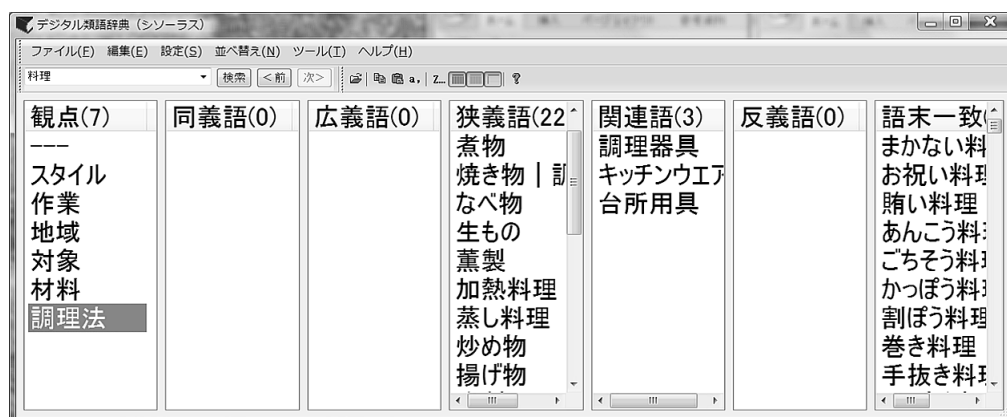


図 4 「調理法」の観点で検索したときの操作画面

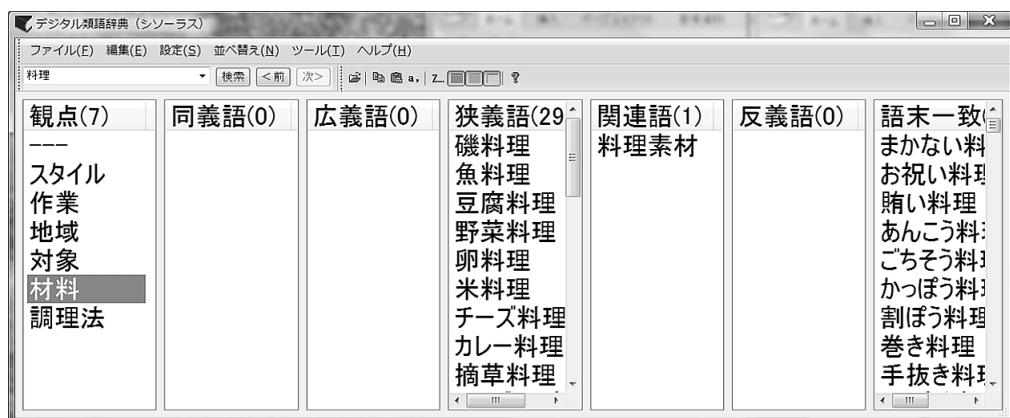


図 5 「材料」の観点で検索したときの操作画面

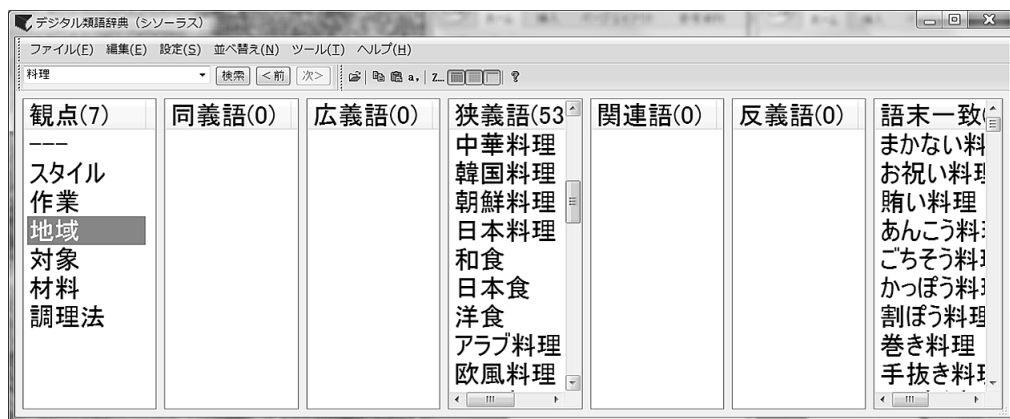


図 6 「地域」の観点で検索したときの操作画面

4.2 未知語の形態素解析

ユーザーがシソーラスに登録されていない用語で検索したときのために、形態素解析をして分解された形態素を自動表示する。画面上の分解された形態素をクリックすると検索できる。

複合語の最後の用語を広義語に、それ以外の用語を関連語とした。接尾辞は同じ意味の自立語に置き換える（下の例を参照）。

例 「未知語解析」を形態素解析すると、下記の3つの用語が表示される。

未知 （関連語）

言語 （広義語） 「語」は接尾辞なので同じ意味の「言語」に変換する。

解析 （広義語）

4.3 語末一致検索

日本語の複合語はほとんどの場合、意味や品詞を決定する用語が語末に、修飾する用語が前方にくる。この性質に着目して語末が同じ用語を取り出すと同じ意味の用語が集められ、狭義語を集めたのと同じような効果を持たせることができる。

例えば「トンボ」をキーにして検索すると、語末が一致として下記の用語が表示される。

狭義語 「アカトンボ」「イトトンボ」「シオカラトンボ」……

ノイズ 「竹トンボ」「尻切れトンボ」「極楽トンボ」

漏れ 「オニヤンマ」「ギンヤンマ」

「トンボ」という言葉を比喩的に用いている場合にノイズになる。

接尾辞

シソーラスには接尾辞も登録できる。

4.4 カスタマイズ機能

ユーザーがどんな用語を関係語として要求するかは個人によって、また置かれた状態によってまちまちである。「非常勤社員」「フリーター」「非正規社員」の同義語である「テンポラリー・ワーカー」などという用語は最近の労働問題を調べているひとには必要であるが、労働問題の歴史を研究しているひとには不要である。用語同士の関係がそのひとの環境、世代で異なることもある。筆者らの世代では、「パソコン」は「コンピューター」の狭義語であるが、最近の社会一般では「パソコン」という言葉の方が一般的になっている。個人別に学習したりする柔軟な機能が必要である。筆者らのシソーラスには次の機能を用意してある。

A. 当面不要な用語をじゃまにならないように、一時隠しておく機能

B. ユーザーが手持ちのシソーラスをファイルから併合する機能

筆者らのシソーラスには専門用語が登録されていない。利用者がそれぞれの専門分野の用語を登録する方式を取っている。ユーザー登録語は優先して表示する。

4.5 エラーレベル・差別語の表示

エラーのある用語と差別語は赤で表示している。また標準でない次のような用語をピンクで表示している。

常用漢字以外を含んでいる用語 例 割烹

表記の揺れ 例 インタフェース

旧地名 例 浦和市

旧機関名 例 文部省

商品名 例 宅急便

画面上正しい用語だけにするために、エラーのある用語と標準でない用語を表示しないよう

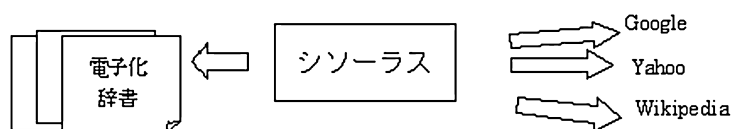


図 7 他のシステムとの接続

にする機能もある。

4.6 インターネット・辞書との接続

「言葉のポータルサイト」を目指している。画面上の用語をクリックすると Google, Wikipedia などのインターネット検索ができる。また同様に電子化された辞書を串刺し検索できるようにしてある。(図 7 参照)

4.7 文章を推敲中の自動検索

ワープロなどで文章を推敲中により適した用語を探すために、クリッピングボード（コピーなどのために切り取った文字列）の文字列でシソーラスを自動的に探す。

5 おわりに

ネット上の記事が現在のペースで増えていくと、キーワードだけの検索ではノイズが多く早晚限界がくると思われる。ノイズを減らすためにも自然文検索のニーズ高まっている。今後日本語解析などを高度化していくためには意味の分野に立ち入らざるを得ないだろう。そのときにシソーラスが多用されるだろう。

「補助的な記述」(1.4) ではすべて同じ (|) で区切って示しているが、それぞれの目的で区切り方を分けておくべきであったと思っている。特に多義語は、それ以外とは意味が異なる。

シソーラス・ファイルのコードは外国語との結合を考えて、unicode を使用している。

これから高度な自然言語処理で活用していくためには、固有名詞も扱わなければならないだろう。

本シソーラスについてご興味のある方は下記までお問い合わせください。

E-mail:kokubu@gengokk.co.jp

謝 辞

本稿の改善に有益なコメントを頂いた査読者の方に感謝いたします。

参考文献

- 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦 (1997). 日本語語彙大系. 岩波書店.
- 医学中央雑誌刊行会編 (2007). 医学用語シソーラス第6版. 医学中央雑誌刊行会.
- 大野晋, 浜西正人 (1985). 類語国語辞典. 角川学芸出版.
- 科学技術振興機構編 (1999). JICST 科学技術用語シソーラス. 科学技術振興機構.
- 言語工学研究所. 「類語.jp」(類語辞書オンライン版). <http://ruigo.jp>.
- 国立国語研究所 (1964). 分類語彙表. 秀英出版.
- 日本工業規格 (1991). シソーラスの構成及びその作成方法: JIS X 0901-1991. 日本規格協会.

付録

A 付録 品詞一覧

カッコ内は活用語尾である.

品詞・活用形	例
名詞	学校
サ変名詞形	勉強 (する)
サ変非名詞形	察 (する)
ザ変	信 (ずる)
一段	生き (る)
カ行五段	書 (く)
カ行五段例外	行 (く)
ガ行五段	泳 (ぐ)
サ行五段	押 (す)
タ行五段	立 (つ)
ナ行五段	死 (ぬ)
バ行五段	遊 (ぶ)
マ行五段	飲 (む)
ラ行五段	走 (る)
ラ行五段例外	おっしゃ (る)
ワア行 5 段	買 (う)
ワ行五段例外	問 (う)

形容詞	青 (い)
形容動詞	閑静 (な)
形容動詞 と／たる形	矍鑠 (たる)
副詞	さっぱり
打ち消しの動詞	年端もいか (ない：助動詞)
打ち消しの形容詞	必要 (ない：形容詞)
連体詞	こんな
接尾辞	*法 見出しに*をつける.

略歴

国分 芳宏 (正会員) : 1966 年東京理科大学理学部応用物理学科卒. 同年日本科学技術情報センター入社. 1985 年株式会社言語工学研究所設立代表取締役就任. 自然言語処理, シソーラス作成に従事. 情報処理学会会員.

岡野 弘行 (正会員) : 1966 年大阪大学理学部物理学科卒. 同年日本科学技術情報センター入社. 2004 年株式会社言語工学研究所入社. シソーラス作成に従事.

(2009 年 2 月 18 日 受付)

(2009 年 6 月 22 日 再受付)

(2009 年 7 月 23 日 採録)